[7] The phrasing is a little foxy here, since I've suddenly switched to talking of understanding sentences instead of talking straightforwardly of sentence truth value. This is because of difficulties with attributing truth value to ambiguous sentences—more of which will be said below. My justification turns on Davidson's claim that one sense of understanding the meaning is knowing the truth conditions. The way one would know the truth conditions using the incorrect (*T* 2) would not allow a proper understanding of what was being said to the child.

[8] This also raises a difficulty for Davidson if the notion of the speaker's changing meaning or intent involves us in irreducible notions of linguistic meaning. See below for remarks on this.

[9] The qualification 'reasonable' is here because there is a good bit of leeway in deciding what is to go into syntax. For example, some of the things Chomsky mentions in [1] as possibilities for inclusion in syntax have generally been considered part of semantics.

[10] There is considerable discussion of this problem in literature in the philosophy of science. I discuss the problem and its effects on reference in "On Criteria of Meaning Change," *British Journal for the Philosophy of Science* 22 (1971), 131–144.

*The Possible in the Actual**

ADAM MORTON

PRINCETON UNIVERSITY

1. We all have beliefs which require quantified modal idioms for their expression. I believe that I could have been the father of six children, if my life had been less lucky. Most of us do not believe that there are possibilia. We do not believe that "I could have been the father of six children" is true by virtue of my or anyone else's paternity of six possible, nonactual, children. It would therefore be disturbing to be told that when one uses such idioms one is referring to possibilia, even though one doubts that there are any. Yet the best analyses we have of the concepts of truth and reference as applied to quantified modal contexts, the various 'possible worlds semantics', can be interpreted as having this disturbing consequence.[1] It is thus of interest to see whether we can find a less 'realistic' attitude towards possible worlds and their inhabitants. Suggestions in this direction have been made by Carnap, Jeffrey, and Quine (see Carnap [1], §41; Jeffrey [4], §12.8; Quine [5]). The purpose of this paper is to present some attitudes and devices which should help to overcome some difficulties of this approach. I first formulate a fairly strict criterion of the ontological

economy of a theory of modality and then describe a way of meeting it, first in informal and then in more formal terms.

I am not going to argue that there are no possibilia. I shall argue just that we are not forced to assume that we refer to any when we use modal idioms. Nor am I going to defend or discuss the intuition that even if there were worlds full of possibilia, facts about them could make no difference to the way things actually are or might have been. But someone who does not share this intuition will not be likely to see much point to following my suggestions.

2. *Ontological Commitment*. To describe the issue more precisely, we must talk about ontological commitment. I am not going to present a theory of ontological commitment, and I am not sure that one is possible, for there seem to be at least two concepts that go by that name. One of them is relational; in terms of it, one may discuss the devices by which languages and theories refer to objects and kinds of objects. Another is non-relational; in terms of it, a nominallist may say that his set-theoretical colleague is committed to sets, although there are none of them for him to be in any way related to. It is the first of these that concerns us now; but I am not going to give a self-sufficient account of it. For our purposes, it is enough to say not what the ontological commitments of a theory are, but what the ontological commitments which a semantical theory ascribes to a theory are.

Suppose that we have an object language *OL* containing various predicates and connectives and also containing operators which bind variables. We may consider not only simple variable binders such as 'for all x' and 'there is an x such that', but also complex ones such as 'for no x' (i.e., 'it is not the case that for some x'), and 'it is possible that there is an x such that', and 'there might have been an x such that'. Let me call all these operators 'quantifiers', and write 'QxP' for the result of applying an arbitrary quantifier binding the variable 'x' to the open sentence 'P'. 'Qx' is not a symbol of *OL*; we use it when describing sentences of *OL*. Suppose that we have a theory θ in a metalanguage for *OL* in which the predicates '$T(s)$' and '$TO(s, x)$', meaning 'closed sentence s of *OL* is true', and 'open sentence s with one free variable is true of object x' respectively, are defined.

Now let us define "'Qx' bears commitment within domain D, according to θ" as: θ has all instances of

$$T(\text{`}QxP\text{'}) \equiv \exists x(x \in D \;\&\; TO(\text{`}P\text{'}, x)$$

as theorems. According to this definition, the existential quantifier of a one-sorted language will bear commitment within the domain of discourse over which the variables range, and '$(x)\sim$' will bear no commitment at all. If *OL* is many-sorted, the quantifications over the different styles of variables will bear commitments within the different domains associated with them. A quantifier will bear commitment within many domains; generally speaking, one is most interested in the smallest domains within which a quantifier bears commitment.

We can now describe the prodigal character of 'realistic' modal semantics more carefully. Such accounts construe the ontological commitments of modal idioms as being different from those of non-modal idioms; modal quantifiers such as 'it is possible that there exists' and 'it is necessary that for all' (or 'there might have been' and 'all *A*s have to be') are credited with bearing commitment to domains to which no non-modal quantifier can bear commitment. It is interesting to note that there is a sense in which it is the universal quantifier rather than the modal operators that is implausibly treated, for according to such accounts, there are sets of objects which the language can be used to discuss but which are not in the range of the word 'all' as used in the language.

It is important to notice that our definition is only applicable relative to an identification of two predicates of the semantical theory in question as being 'is true' and 'is true of'. If the intention of the semantical theory is just to define the concepts of validity, consequence, and consistency, then it may well involve neither 'true' nor 'true of'. Instead, it will use the relativized predicates 's is true under interpretation I' and 's is true of x under interpretation I'. However, truth cannot be defined in terms of truth-under-I without picking some interpretation as the intended one. And if one is to pick an intended interpretation, one must do so in a way that treats predicates and quantifiers plausibly.

For example, a typical modal semantics which is only concerned with validity, consequence, and consistency may be a theory of structures $\langle W, A, R, F\rangle$, where W is a set of 'possible worlds', A is a member of W, R is a relation between the members of W, and F is a function which assigns an extension in each member of W to each predicate in the language. One can then in familiar ways define 's is true under $\langle W, A, R, F\rangle$' for variable $\langle W, A, R, F\rangle$. The Ws may be any collections of sets at all; there is no need to interpret any of them as containing sets of possibilia. But if one is to define 's is true', one must pick some W to be the set of possible worlds, A to be the actual world, R to be the accessibility relation, and F to assign to each predicate its extension. If by 'assign to each predicate its extension' we mean 'assign to each predicate a set, for each possible world, such that that predicate is true of the members of that set', then our theory represents the language as being commited to non-actual entities, since for a sentence such as '$(x) \sim \Diamond P(x) \& \Diamond \exists x P(x)$' to be true, it is necessary that there be some member of some world accessible from A which is not in A and of which 'P' is true. (The quantifier here is '$(x) \sim \Diamond$— & $\Diamond \exists x$—'.) One could avoid this result by not interpreting the relation between 'P' and the member of the non-actual world as being 'is true of'. But it is not at all clear how this could be done. One cannot simply leave it undefined, at least if one is trying to give truth conditions in a philosophically informative way.

3. *Simulation.*

> Wimsey waited at the window till he saw Dalziel leave the next-door cottage and carefully lock the door behind him, replacing the key in its hiding-place. When the hum of the car had died away in the distance, he ran hastily down the stairs, and across to the garage.
>
> "Corpse!" he cried.
>
> "Yessir!" said the corpse, smartly.
>
> "While that ghastly blighter was nosing around, I—in my role as murderer, you understand—had an awful thought. All this time you're getting stiff. If I leave you like that I shall never be able to pack you into the back of the car. Come out, sir, and be arranged in a nice hunched-up position."
>
> "Don't you dump me in the car earlier?"
>
> "No, or you wouldn't look natural. I lay you out on the floor to set." ([6]: 261)

Often in detective novels the detective shows that the crime could have been committed in a certain way by elaborately reconstructing the crime as thus committed, often playing the murderer himself. There are immense differences between the real crime, or the crime as supposed by the detective, and the reconstruction. For one thing, no one is killed, and thus the corpse can speak. For another, the participants in the reconstruction need not have participated in the crime, and, if they have, they need not play themselves. They may have only very slight resemblances to the individuals they play. No live person can be very much like a corpse.

By such a piece of play acting, I can show that something could have existed which does not exist. Thus, I can show how there might have been a footprint in the flowerbed yesterday, although there was not, by putting one there today.

Devices for demonstrating the possibility of processes and outcomes need not exhibit even the small degree of similarily between an individual and the role he plays that is found in the detective story examples. One can, for example, show how things could have turned out by means of a computer simulation; then electrical impulses play the roles of masses of people or of planetary bodies. Presumably, the connection between possibility and conceivability (complex though it is, neither being contained in the other) holds because the human nervous system can operate as an analog machine. (However, analog computation is in general concerned with considerably looser resemblances between model and situation modelled than will be dealt with here.)

Still, a demonstration or simulation is only a device for showing how things might have been. There are ways the world could have been which we cannot act out or even describe, one reason being that some of these ways are beyond our imagination. What makes it *true* that things could have been so-and-so is not that someone happens to make an appropriate simulation, but rather that if a certain simulation had been carried out, it would have had a certain import. How can this be spelled out in a non-circular way?

4. *Partial Worlds*. One might at this point try to construe actual simulations of possible situations, or, more likely, set-theoretical constructions representing simulations that could be made, as ersatz possible worlds. A proposition would then be possible when there is a simulation in which it holds.

One difficulty this approach would run into derives from the fact that one cannot simulate a situation in which more individuals exist than there actually are. Thus, sentences such as

$$(*)\ \exists x \exists y (z)(z = x \vee z = y) \supset \Box \exists x \exists y (z)(z = x \vee z = y)$$

will turn out to be valid. Yet clearly even if there are only n individuals, there could have been $n + 1$; at any rate we should not decide the opposite as a matter of logic.

One can brazen one's way out of this difficulty. However, a more serious philosophical objection remains. It is in fact the strongest argument for taking a realistic attitude towards possible worlds and their inhabitants. It arises when one tries to account for iterated modal idioms. One seems then to be forced to make distinctions between possible worlds that will in general not be expressible just in terms of the non-modal vocabulary of the language one is studying. Non-modal truth does not seem to determine modal truth. For example, I cannot lift two hundred pounds, but there surely are worlds in which I am stronger than I actually am and thus can lift two hundred pounds, even though in some of these worlds I live my life in exactly the way I actually do and never lift two hundred pounds. Thus, if s_1, s_2,... are all the non-modal sentences in a given vocabulary which are actually true, then it is possible that (s_1 and s_2 and... and it is possible that q), although actually s_1 and s_2 and ... and it is not possible that q. Now surely such worlds do differ in some non-modal respect from the real world, in that some feature of each such world (more androgen in my bloodstream, perhaps) is responsible in that world for my great strength. But the relevant features need not be describable in terms of the predicates available in any language under consideration and may indeed be beyond human capacity to describe. Thus, the objects in non-actual worlds must have properties beyond our ability to make true predications of them, just as actual objects have, and how can this be unless there really are such objects and they really have properties in the way that actual objects do?

This difficulty forces us to take account of the way that the properties of objects can exceed our ability to describe them in terms of any given language. However, it seems possible to make our analysis turn on just this characteristic of the actual world without having to assume that it holds of any non-actual worlds. In fact, I think that we are now approaching an essential feature of modal idioms. We use 'possibly', 'necessarily', 'might', and 'would' to indicate features of the world which we cannot or do not want to describe explicitly. Moreover, we use these words to indicate features of the world which in some way underlie and are responsible for those features which we can describe explicitly.

This way of putting things leads, I think, to a way out of our difficulties. For to get the effect of the last paragraph, we do not need to use the notion of a possible world as it is usually understood, that is, of a world in which some things are true which are actually false. What we need are *partial worlds*. A partial world is a partial specification of the real world. In it, nothing is true which is actually false, but some things may be neither true nor false which are actually true (or false). A partial world is given by specifying the extensions of some, but not necessarily all, of the properties of members of the universe of discourse. Some properties occur in more of the partial worlds than others; these are the more fundamental ones. A sentence is possible if its truth conditions are in all fundamental respects like those of a sentence which is true. That is to say, the extensions of the predicates from which it is constructed are in some partial world indistinguishable from those of some sentence which is actually true.

This rough formulation, which will be improved on in the next section, already allows us to go back and see why the notion of a simulation was attractive. When a simulation is successful, it shows that two predicates (or two objects) are fundamentally the same, so that what is true in terms of the one might have been true in terms of the other.

Two metaphors may be useful. First, think of a partial world as an embryo world. Think of the world as developing its features successively, starting from the more fundamental features and progressing to the more superficial. A sentence is then possible if it is not ruled out until a relatively late stage of the world's gestation, if at all. Second, think of a partial world as a fuzzy photograph of the real world. A sentence is then possible if its falsity only shows up, if at all, on fairly sharp photographs of the world. But metaphors are no substitute for a definition. The next section gives one.

5. *A Semantical Theory*. At this point, we could simply describe a reinterpretation of a standard possible-worlds semantics, so that the array of possible worlds is taken as describing an array of partial worlds. It would be clumsy, however, and artificial. We stand to learn more by making a semantical theory which directly reflects the idea of a partial world.

Consider a modal object language *OL* whose vocabulary consists of: a single two-place relation 'R', plus '$=$' and the logical constants '&', '$\sim$', '$\exists$', '$\Diamond$', '$\rightarrow$', '$\mathscr{M}$', plus variables and brackets. '$\rightarrow$' is a subjunctive conditional; '$p \rightarrow q$' means 'if p were the case, then q would be'. '$\mathscr{M}$' is a subjunctive modal operator, meaning 'might'. '$\vee$', '$\supset$', '( )', '$\Box$', '$\dashv$' are defined in familiar ways. We are going to define truth and satisfaction for *OL*, in a metalanguage which contains 'R' and the non-modal logical terms, but in which the operators '$\Diamond$', '$\rightarrow$', and '$\mathscr{M}$' do not occur. The truth definition will proceed along roughly model theoretical lines, as is usual, for we are interested in characterizing the concepts of validity and logical consequence for *OL*, as well as capturing the Tarski biconditionals (relative to a translation into the metalanguage).

We begin by choosing a set D as the domain over which the variables of *OL* are to range. It is most convenient next to expand *OL* by adding a number of additional predicate symbols $P_1, \ldots, P_n$, to obtain a language OL^+. These symbols are going to represent properties of objects in D which are not explicitly mentioned in *OL*. We next define a structure @, called the *total world*, as the ordered $(n + 2)$-tuple $\langle D, \Pi_0, \Pi_1, \ldots, \Pi_n \rangle$, where Π_0 and the other Π_i are relations over D, intended as the extensions of R and the P_i, respectively. Π_0 is of course dyadic; the other Π_i could be of any adicities, but for simplicity I shall assume that they are all monadic. We also need *partial worlds*, which are tuples $\langle \Delta, \Pi_{\alpha_1}, \ldots, \Pi_{\alpha_m} \rangle$ obtained by selecting a subset Δ of D and some selection of the Π_i (restricted to Δ). We say that each Π_{α_i} *occurs in* the partial world. Note that a partial world differs from @ only in that it need not contain all members of D, and some of the Π_i may not occur in it. It assigns to no predicate a wrong or extended extension.

With the system of worlds we associate a partial ordering, i.e., a transitive, reflexive, and antisymmetrical relation over a certain set (whose members may be anything at all, for our purposes), which has a unique minimal element 0, and which has the worlds as maximal elements. Moreover, each world is connected to 0 by a single path. The path connecting @ to 0 plays a special role; call it $P_@$. The whole structure can be visualized as a tree, with 0 at the base of the trunk and the worlds at the tips of the branches. $P_@$ is the trunk, which continues to the very top.

World W' is *accessible from* world W if the path leading to W meets $P_@$ at a point no higher than that at which the path leading to W does, and no path branches from $P_@$ between these points. One world is thus accessible from another when it can be reached by 'unravelling' the other a little (but not too much) below the point at which the path of the other meets $P_@$. Note the special role that $P_@$ plays in the definition of accessibility. It is best not to include 0 in $P_@$, in order to allow there to be worlds from which @ is not accessible. It would have been possible to define the ordering relation in terms of the degree of approximation of the partial worlds to @.

Five considerations motivate the formal definitions of satisfaction and truth:

(a) Since not all the predicates of OL^+ occur in each partial world, the truth value of some sentences in some partial worlds may not be determined in the usual way. But we can find truth values for some of these sentences in some partial worlds in an unusual way. The idea is that although a sentence may not have a truth value in a partial world by the straightforward reckoning, its truth conditions in that world may be 'just like' those of some sentence which is actually true. The partial world then exhibits no reason why the truth values of the sentences should be different. By 'just like' I mean the following: the extensions of the predicates involved in the two sentences are indistinguishable in the partial world in question. I give a precise definition below.

(b) "$\Diamond A$" is true in a partial world W if there is a partial world accessible from W in which "A" is 'just like' a sentence which is actually true.

(c) "$A \rightarrow B$" is true in a partial world W if "B" is 'just like' a true sentence in all worlds which branch from W at as high as possible a point consistent with "A" being 'just like' a true sentence in them. This definition is similar to, and in imitation of, Stalnaker's. (See Stalnaker [7].) It differs from his mostly in that it allows for the possibility that there may be several worlds which are equally similar to the actual one, in which "A" is true, and requires, as seems right, that "B" should be true in all of them. (David Lewis tells me that this modification of Stalnaker's idea has occurred to a number of people, including himself and Stalnaker.)

(d) "$\mathscr{M}A$" is true in a partial world W if there is a partial world W' such that the paths of W, W', and @ have a common point, and "A" is 'just like' a true sentence in W'.

(e) '$=$' requires some complication of the simple model described under (a). A sentence involving '$=$' is not taken to be 'just like' another sentence got by substituting predicates for indistinguishable predicates, but rather to be 'just like' a set of sentences got by making all possible substitutions for '$x_n = x_m$' of predicates of the form '$P_i(x_n) \equiv P_j(x_m)$', where P_i and P_j are predicates occurring in the partial world in question (and otherwise substituting predicates for predicates as described in (a)). Identity is taken as being 'just like' indistinguishability.

To state this formally, let s, s',... be sequences $\langle s_1, s_2, \ldots\rangle$ of members of D. We will define three predicates, 's satisfies "A" in W', 'Φ is a simulation', and 's satisfies "A" in W under Φ', by simultaneous recursion. The second of these can be defined just in terms of non-modal satisfaction in a partial world, and it is simplest to give this definition first.

If D_1 and D_2 are subsets of D, then D_1 and D_2 are *indistinguishable in W* if every non-modal open sentence of OL^+ constructed from P_i, whose Π_i occur in W, and which is satisfied by every member of D_1 is satisfied by every member of D_2, and vice versa. Similarly for sets of ordered pairs of members of D, etc. If d_1 and d_2 are members of D, then d_1 and d_2 are indistinguishable in W if their unit sets are indistinguishable in W.

Let 'Φ is a simulation in W' be defined as the disjunction of (A) and (B), as follows.

(A) Φ is a function which takes sentences in OL^+ to sentences in OL^+ obtained as follows. First replace any predicates P_i by predicates Q_i (the Q_i may be complex), provided that each P_i does not occur in W, and the intersection of Π_i and the domain of W is indistinguishable in W from the class of objects satisfying Q_i in @. Then replace each occurrence of '$x_m = x_n$' by the conjunction of all '$P_j(x_m) \equiv P_j(x_n)$', where Π_j occurs in W.[2]

(B) Φ is a function which takes members of the domain of W to other members of the domain of W, such that $\Phi(x)$ is always indistinguishable in W from x. By a $\Phi(s)$, where s is a sequence $\langle s_1, s_2, \ldots \rangle$, I mean a sequence $\langle s_1', s_2', \ldots \rangle$ where each s_i' is either s_i or $\Phi(s_i)$.

We define 's satisfies "A" in W' and 's satisfies "A" in W under Φ' by:

(i) s satisfies "$R(x_n, x_m)$" in W if and only if s_n and s_m are in the domain of W, Π_0 occurs in W, and $\langle s_n, s_m \rangle \in \Pi_0$.

(ii) s satisfies "$x_n = x_m$" in W if and only if s_n and s_m are in the domain of W and $s_n = s_m$.

(iii) s satisfies "A" in W under Φ if and only if Φ is a simulation in W and either some $\Phi(s)$ satisfies Φ("A") in @ or s satisfies "A" in W.

(iv) s satisfies "$\Diamond A$" in W if and only if there is a partial world W' which is accessible from W and a simulation Φ in W' such that s satisfies "A" in W' under Φ.

(v) s satisfies "$A \rightarrow B$" in W if and only if for all worlds W' and simulations Φ in W' if s satisfies "A" in W under Φ and there is no world W'' and simulation Φ' in W'' such that the path of W'' meets that of W at a later point than that of W' does, and s satisfies "A" in W' under Φ, then s satisfies "B" in W' under Φ.

(vi) s satisfies "$\mathscr{M}A$" in W if and only if there is a partial world W' whose path and that of W have a common point on $P_{@}$, and a simulation Φ in W', such that s satisfies "A" in W' under Φ.

(vii) The clauses for '$\sim$', '$\exists$', '&', both for satisfaction in W and for satisfaction in W under Φ, are completely standard.

We may now define '$T(s)$' as 's is a closed sentence of OL, and all sequences satisfy s in @', and '$TO(s, x)$' as 's has one free variable, and for all n, if that free variable is the n-th variable, then all sequences whose n-th member is x satisfy s.'

These definitions have the following consequences:

(a) All quantifiers in OL range over D. No modal quantifier bears commitment within a domain within which no non-modal quantifier bears commitment.

(b) Although the domain of each partial world is a subset of the set D of actual objects, sentences such as (*), mentioned in section 4, are not valid. (For an exception, see the next paragraph.) To construct a counterexample to (*), take D to be $\{a, b\}$. "$\Diamond\exists x\exists y\exists z(\sim(x = y)\ \&\ \sim(y = z)\ \&\ \sim(x = z))$" can be true. For suppose that the extension of R is $\{\langle a, b\rangle\}$, and that in some partial world accessible from @, a is indistinguishable from b. Let us abbreviate '$\langle a, b, \ldots\rangle$ satisfies "$A(x_1, x_2)$"' by '"$A(a, b)$" is true'. Now "$R(a, b)\ \&\ R(a, b)\ \&\ \sim R(b, a)\ \&\ \sim R(b, a)$" is true in @. So "$\exists z(R(a, b)\ \&\ R(a, z)\ \&\ \sim R(b, a)\ \&\ \sim R(z, a))$" is true in @, so "$\exists z(R(a, b)\ \&\ R(b, z)\ \&\ \sim R(b, a)\ \&\ \sim R(z, b))$" is true in W under the appropriate simulation (which simulates a with b). Therefore "$\exists x\exists y\exists z(R(x, y)\ \&\ R(y, z)\ \&\ \sim R(y, x)\ \&\ \sim R(z, x))$" is true in W under this simulation. But this suffices to make x, y, and z distinguishable, so $\exists x\exists y\exists z(\sim(x = y)\ \&\ \sim(y = z)\ \&\ \sim(z = x))$ is true in W under this simulation. The intuitive pattern of the argument is as follows. The world contains just a and b. a bears R to b, but not vice versa. But since a and b are indistinguishable (in W), b might too have borne R to something and not vice versa. This something could not be b, since b bore R to it and not vice versa, and it could not be a, since b bore R to it and b does not bear R to a. Therefore, there could have been a third object.

However, there is, as far as I can see, no way of making "$\Diamond\exists x\exists y(\sim(x = y))$" true in a domain of one individual. This leads to counterintuitive results, and, as a result, I think we should restrict our attention to domains of at least two individuals, just as in standard quantification theory one restricts one's attention to domains of at least one individual.

(c) By imposing conditions on the path structure, we can make the accessibility relation transitive or symmetrical or both. We can therefore get as valid sentences under these conditions all instances of the theorems of the propositional forms of S4, S5, and the 'Brouwerian' system. It is less clear to me what sorts of quantified modal principles come out valid, and under what conditions.

(d) *OL* did not contain any proper names, but they would have presented no difficulty. Each name would simply have a denotation, and it would denote it in every partial world in which the denotation appears. No trans-world identification problems arise. However, it is not consistent with our scheme that a proper name, say 'Pegasus', should name a non-actual object. This seems right to me.

Something similar happens with predicates. If 'unicorn' is taken as being atomic, then the semantics above has as a consequence that "$\Diamond \exists x$ Unicorn (x)" is false. But if 'unicorn' is taken as an abbreviation for, say, 'horse with a horn', then the sentence has a chance of being true.

6. *Objections and Replies.*

Objection. There is no point to satisfying a constraint on ontological commitment which can be satisfied by a theory which appeals to types of objects, such as partial worlds and paths connecting them, that are just as far beyond the range of object-language quantifiers as possibilia are.

Reply. One always needs more in the metalanguage. That seems to be a fact of life. But one need not represent the use of the metalanguage as constituting reference to any of the metalinguistic apparatus, and we do not.

Objection. The whole subject is a red herring. Talk of possible worlds does nothing to tell us what is possible or what 'possible' means.

Reply. There is something to this. The most that enterprises such as this one can do is to clarify the logical properties of modal sentences. To do so it is not necessary to say what is in fact possible. Deciding what is possible is in any case not something that philosophers can do in their studies, for it is often an empirical discovery that a process or event is possible. But saying what 'possible' means *is* something that philosophers should try to do, and the largest part of the job is left untouched by considerations such as have filled these pages. However, logical considerations can be useful in three ways. In the first place, they can reduce the problem of explaining an idiom whose logical form is problematic to one of describing the extension of various relations, in our case the relation which defines the path structure. In the second place, they can help in discovering connections between different idioms and common features of them. The apparatus I have described allows us to give a unified treatment of 'possibly', 'might', and the counterfactual 'would', and I think that some similar apparatus is inevitable if we are to understand the whole range of modal idioms available in a language like English. And in the third place, logical considerations can impose constraints on a theory of meaning; an explanation of modal idioms should consist in an explanation of the terms left undefined by some formal semantics. Such constraints are only welcome if the formal semantics embodies plausible prejudices. The account given here embodies two prejudices. One is about reference: when talking about how the world could have been, or might have been, one is talking just about the world. The other is more metaphysical: we use modal idioms to indicate, without describing, the secret powers of nature. I have not argued for either of these prejudices. However, I have tried to show that they meet one of the more basic requirements on an attitude towards modality, that it be capable of motivating a formal account of validity and truth.

References

[1] Carnap, R., *Meaning and Necessity*, second edition (Chicago: University of Chicago Press, 1956).
[2] Goodman, N., *Fact, Fiction, and Forecast* (Cambridge, Mass.: Harvard University Press, 1955).
[3] Hughes, G. and Cresswell, M., *An Introduction to Modal Logic* (London: Methuen, 1968).
[4] Jeffrey, R., *The Logic of Decision* (New York: McGraw-Hill, 1965).
[5] Quine, W. V., "Propositional Objects," in *Ontological Relativity and Other Essays* (New York: Columbia University Press, 1969).
[6] Sayers, D., *Five Red Herrings* (New York: Avon, 1968).
[7] Stalnaker, R., "A Theory of Conditionals," in N. Rescher (ed.), *Studies in Logical Theory* (Oxford: Blackwell, 1968).

Notes

* Fabrizio Mondadori has given me a lot of help when writing this. Russell Trenholme and the referee made helpful suggestions on earlier drafts. My debt to Chapter II of Goodman [2] should be obvious.

[1] For a description of the standard modal systems and semantics, and an adequate bibliography, see Hughes and Cresswell [3].

[2] We can use conjunction here because the number of Π_i is finite. In the more general case in which the set of Π_i is the power set of D, the definition would have to be more complicated.

Opacity and the Ought-to-Be

LOU GOBLE

UNIVERSITY OF NORTH CAROLINA, CHAPEL HILL

It ought to be that 'it ought to be that...' is referentially transparent, but it isn't. (This shows that even in logic this may be less than the best of all worlds.)

When one considers ordinary arguments involving deontic contexts, like 'it ought to be that...', inferences by replacement of identically referring terms, e.g.:

(1) The man who robbed the bank ought to be punished.
(2) Jones = the man who robbed the bank.
Therefore,
(3) Jones ought to be punished,

all seem to be valid, as do inferences by Existential Generalization, e.g., from (1) or (3) to

(4) There is someone who ought to be punished.

In this regard, deontic logic would seem to be the nicest member of the modal logic family.

This is how it should be. One expects no problems interpreting quantification into deontic contexts. Unlike the logics for alethic necessity or belief, with deontic logic there is not even a temptation to think that the values of the individual variables are intensions, or individual concepts, or possibilia (permissibilia ?). Statements of obligation, permission, prohibition, etc. seem to be directly about persons or things regardless of how they are conceived or described. The deontic operators thus certainly seem to be extensional; nevertheless, when principles of extensionality are assumed for standard systems of deontic logic, paradox appears.[1]